

**Computer Learning for Computer Vision**

Code: 43083  
ECTS Credits: 6

Degree	Type	Year	Semester
4314099 Computer Vision	OB	0	1

**Contact**

Name: Ramón Baldrich Caselles  
Email: Ramon.Baldrich@uab.cat

**Use of Languages**

Principal working language: english (eng)

**Teachers**

Ernest Valveny Llobet  
Maria Vanrell Martorell  
Juan Carlos Moure Lopez  
Fernando Luis Vilariño Freire  
Dimosthenis Karatzas  
Pau Rodriguez Lopez  
Lluís Gomez Bigorda

**Prerequisites**

- Degree in Engineering, Maths, Physics or similar.
- Programming Skills in Python.

**Objectives and Contextualisation**

The objective of this module is to introduce the Machine learning techniques for solving computer vision problems. Machine learning deals with the automatic analysis of large scale data. Nowadays it conforms the basics of many computer vision methods, specially those related to visual pattern recognition or classification, where 'patterns' encompasses images of world objects, scenes and video sequences of human actions, to name a few.

This module presents the foundations and most important techniques for the classification of visual patterns, mainly focusing on supervised methods. Also, related topics like image descriptors and dimensionality reduction are addressed. As much as possible, all these techniques are tried and assessed on a practical project concerning scene description from pictures, together with the standard metrics and procedures for performance evaluation like precision-recall curves and k-fold cross-validation.

The learning outcomes are:

- (a) Distinguish the main types of ML techniques for computer vision: supervised vs. unsupervised, generative vs. discriminative, original feature space vs. feature vector kernelization.

(b) Know the strong and weak points of the different methods, in part learned while solving a real pattern classification problem.

(c) Being able to use existing method implementations and build them from scratch.

The module goes in depth in two main approaches to introduce ML into the image classification problem. Using: a) handcrafted image description, b) data driven image description. On the first case the Bag of Words is used, on the second one, the Deep Learning approach. The DL content is developed extensively providing both, theoretical basis of the different parts of modern Neural Networks architectures, and best practices to apply it on real applications.

## Competences

- Accept responsibilities for information and knowledge management.
- Choose the most suitable software tools and training sets for developing solutions to problems in computer vision.
- Conceptualise alternatives to complex solutions for vision problems and create prototypes to show the validity of the system proposed.
- Continue the learning process, to a large extent autonomously.
- Identify concepts and apply the most appropriate fundamental techniques for solving basic problems in computer vision.
- Plan, develop, evaluate and manage solutions for projects in the different areas of computer vision.
- Solve problems in new or little-known situations within broader (or multidisciplinary) contexts related to the field of study.
- Understand, analyse and synthesise advanced knowledge in the area, and put forward innovative ideas.
- Use acquired knowledge as a basis for originality in the application of ideas, often in a research context.
- Work in multidisciplinary teams.

## Learning Outcomes

1. Accept responsibilities for information and knowledge management.
2. Choose computational learning techniques and train them to resolve a particular project.
3. Continue the learning process, to a large extent autonomously.
4. Identify the basic algorithms of computational learning and their application.
5. Identify the best representations that can be defined for solving computational learning problems.
6. Solve problems in new or little-known situations within broader (or multidisciplinary) contexts related to the field of study.
7. Understand, analyse and synthesise advanced knowledge in the area, and put forward innovative ideas.
8. Use acquired knowledge as a basis for originality in the application of ideas, often in a research context.
9. Use computational learning techniques to plan, develop, evaluate and manage a solution to a particular problem.
10. Work in multidisciplinary teams.

## Content

1. Introduction to machine learning
2. Experimental Setup
3. Embeddings: SVM and Random Forest
4. Introduction to Neural Networks
5. Introduction to Deep Learning
6. Convolutional Neural Networks
7. Training: data pre-processing, initialization, gradient optimization
8. Image Classification

9. Understanding and visualizing CNNs
10. Efficient methods for Deep Learning

## Methodology

Supervised sessions:

- Lecture Sessions, where the lecturers will explain general contents about the topics. Some of them will be used to solve the problems.

Directed sessions:

- Project Sessions, where the problems and goals of the projects will be presented and discussed, students will interact with the project coordinator about problems and ideas on solving the project (approx. 1 hour/week)
- Presentation Session, where the students give an oral presentation about how they have solved the project and a demo of the results.
- Exam Session, where the students are evaluated individually. Knowledge achievements and problem-solving skills

Autonomous work:

- Student will autonomously study and work with the materials derived from the lectures.
- Student will work in groups to solve the problems of the projects with deliverables:
  - Code
  - Reports
  - Oral presentations

## Activities

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Lecture sessions	20	0.8	4, 5, 8
Type: Supervised			
Project follow-up sessions	8	0.32	1, 7, 4, 5, 6, 3, 2, 8, 10, 9
Type: Autonomous			
Homework	113	4.52	1, 7, 4, 5, 6, 3, 2, 8, 10, 9

## Assessment

The final marks for this module will be computed with the following formula:

$$\text{Final Mark} = 0.4 \times \text{Exam} + 0.55 \times \text{Project} + 0.05 \times \text{Attendance}$$

where,

Exam: is the mark obtained in the Module Exam (must be  $\geq 3$ )

Attendance: is the mark derived from the control of attendance at lectures (minimum 70%)

Projects: is the mark provided by the project coordinator based on the weekly follow-up of the project and deliverables. All accordingly with specific criteria such as:

- Participation in discussion sessions and in team work (inter-member evaluations)
- Delivery of mandatory and optional exercises.
- Code development (style, comments, etc.)
- Report (justification of the decisions in your project development)
- Presentation (Talk and demonstrations on your project)

## Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Exam	0.4	2.5	0.1	1, 7, 5, 6, 2, 8, 9
Project	0.55	6	0.24	1, 7, 4, 5, 6, 3, 2, 8, 10, 9
Session attendance	0.05	0.5	0.02	1, 4, 5, 6, 8

## Bibliography

Journal papers:

1. Barber, D. "*Bayesian Reasoning and Machine Learning*". Cambridge University Press, 2012.
2. Yoshua Bengio. "*Learning Deep Architectures for AI*". Foundations and Trends in Machine Learning, Vol. 2, No. 1, 2009.
3. Christopher J. C. Burges. "*Dimension Reduction: A Guided Tour*". Foundations and Trends in Machine Learning, Vol. 2, No. 4, 2009.
4. Christoph H. Lampert. "*Kernel Methods in Computer Vision*". Foundations and Trends in Computer Graphics and Vision, Vol. 4, No. 3, 2008.
5. Tinne Tuytelaars and Krystian Mikolajczyk. "*Local Invariant Feature Detectors: A Survey*". Foundations and Trends in Computer Graphics and Vision, Vol. 3, No. 3, 2007.

Books:

1. Ian Goodfellow, Yoshua Bengio and Aaron Courville. "*Deep Learning*". 2016. Cambridge, MA, USA: The MIT Press. ISBN: 978-0262035613
2. Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar, "*Foundations of Machine Learning*" MIT Press, 2012. <http://www.cs.nyu.edu/~mohri/mlbook/>
3. Z.H. Zhou. "*Ensemble Methods: Foundations and Algorithms*". Chapman & Hall/CRC, 2012.

Reports:

1. Criminisi, A. and Shotton, J. and Konukoglu, E. "*Decision Forests for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning*". Technical report MSR-TR-2011-114. Microsoft Research, 2011.  
[http://research.microsoft.com/pubs/155552/decisionForests\\_MSR\\_TR\\_2011\\_114.pdf](http://research.microsoft.com/pubs/155552/decisionForests_MSR_TR_2011_114.pdf)